

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 04:16:36 ; Search time 2344.34 Seconds  
(without alignments)  
11407.790 Million cell updates/sec

Title: US-09-403-262a-1  
Perfect score: 1729  
Sequence: 1 cctctgctctcccccag.....cttattttatttaatta 1729

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_inv3:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vl:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vl1:\*  
59: gb\_vl2:\*  
60: gb\_vl3:\*  
61: gb\_vl4:\*  
62: gb\_vl5:\*  
63: gb\_vl6:\*  
64: gb\_vl7:\*  
65: gb\_vl8:\*  
66: gb\_vl9:\*  
67: gb\_vl10:\*  
68: gb\_vl11:\*  
69: gb\_vl12:\*  
70: gb\_vl13:\*  
71: gb\_vl14:\*  
72: gb\_vl15:\*  
73: gb\_vl16:\*  
74: gb\_vl17:\*  
75: gb\_vl18:\*  
76: gb\_vl19:\*  
77: gb\_vl20:\*  
78: gb\_vl21:\*  
79: gb\_vl22:\*  
80: gb\_vl23:\*  
81: gb\_vl24:\*  
82: gb\_vl25:\*  
83: gb\_vl26:\*  
84: gb\_vl27:\*  
85: gb\_vl28:\*  
86: gb\_vl29:\*  
87: gb\_vl30:\*  
88: gb\_vl31:\*  
89: gb\_vl32:\*  
90: gb\_vl33:\*  
91: gb\_vl34:\*  
92: gb\_vl35:\*  
93: gb\_vl36:\*  
94: gb\_vl37:\*  
95: gb\_vl38:\*  
96: gb\_vl39:\*  
97: gb\_vl40:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1729	100.0	1729	9	A84072	A84072 Sequence 1
2	1729	100.0	1729	12	AF098674	AF098674 Lycopersi
3	1729	100.0	57047	14	LES303345	AJ303345 Lycopersi
4	1223.2	70.7	1296	9	A84080	A84080 Sequence 9
5	294.2	17.0	87967	12	AC005223	AC005223 Arabidops
6	294.2	17.0	109171	12	AC002328	AC002328 Genomic s
7	275.2	15.9	27056	14	CRU303349	AJ303349 Capsella
8	218.8	12.7	687	9	A84084	A84084 Sequence 13

9	141.6	8.2	2210	13	ATFGAL	Y11336 Arabidopsis
10	140	8.1	2201	14	ATY15194	Y15194 Arabidopsis
11	140	8.1	95137	12	AC005560	AC005560 Arabidops
12	133.6	7.7	1951	13	ATFRGA2	Y11337 A. thaliana
13	133.6	7.7	1964	9	A64697	A64697 Sequence 1
14	133.6	7.7	1964	14	ATY15193	Y15193 Arabidops
15	133.6	7.7	132699	12	AC006917	AC006917 Genomic s
16	132	7.6	83289	13	ATAC009895	AC009895 Arabidops
17	125.4	7.3	1779	10	AX081276	AX081276 Sequence
18	123.8	7.2	1779	10	AX081278	AX081278 Sequence
19	115.4	6.7	1642	9	A64705	A64705 Sequence 9
20	115.4	6.7	1643	9	A64699	A64699 Sequence 3
21	105.8	6.1	1359	12	AF036300	AF036300 Arabidops
22	105.8	6.1	2130	13	AF210731	AF210731 Arabidops
23	105.8	6.1	4081	13	ATA224957	AJ224957 Arabidops
24	105.8	6.1	85702	12	AC020665	AC020665 Arabidops
25	105.8	6.1	108365	12	AC015447	AC015447 Arabidops
26	105.4	6.1	1642	9	A64701	A64701 Sequence 5
27	104.8	6.1	69850	12	AB010700	AB010700 Arabidops
28	104.2	6.0	26604	13	ATK3M16	AL391130 Arabidops
29	102.4	5.9	100512	13	AC010795	AC010795 Arabidops
30	100.6	5.8	92745	12	AC079279	AC079279 Arabidops
31	100.6	5.8	105733	12	AC012561	AC012561 Arabidops
32	98.4	5.7	84196	14	ATFR3A5	AL132979 Arabidops
33	92.6	5.3	1636	9	A64703	A64703 Sequence 7
34	92.2	5.3	1065	12	AF036302	AF036302 Arabidops
35	86.2	5.0	1793	12	AF133443	AF133443 Arabidops
36	86.2	5.0	59372	12	AB023039	AB023039 Arabidops
37	84.8	4.9	1078	12	AF036308	AF036308 Arabidops
38	84.6	4.9	1231	12	AF036301	AF036301 Arabidops
39	82.4	4.8	104830	12	AC022464	AC022464 Genomic s
40	81.8	4.7	83689	12	AB017067	AB017067 Arabidops
41	80	4.6	1634	13	AF210732	AF210732 Arabidops
42	80	4.6	96654	12	AC007211	AC007211 Arabidops
43	77.4	4.5	3663	12	AB012633	AB012633 Arabidops
44	76	4.4	198788	13	ATCHRIV46	AB126339 Nicotiana
45	74.4	4.3	7218	10	166494	166494 Sequence 14

ALIGNMENTS			
RESULT	1		
LOCUS	A84072		
DEFINITION	Sequence 1 from Patent WO9846759.		
ACCESSION	A84072		
VERSION	A84072.1	GI:6733213	
KEYWORDS			
ORGANISM	tomato.		
SOURCE	Lycopersicon esculentum		
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasit; Solanales; Solanaceae; Solanum; Lycopersicon.		
AUTHORS	1 (bases 1 to 1729)		
TITLE	Theres, N.		
JOURNAL	PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED ABSCISSION AREA FORMATION		
FEATURES	Patent: WO 9846759-A 1 22-OCT-1998;		
SOURCE	Theres NIKOLAUS (DE)		
ORIGIN	Location/Qualifiers		
	1..1729		
	/organism="Lycopersicon esculentum"		
	/db_xref="taxon:4081"		
BASE COUNT	487 a	372 c	294 g 576 t
Query Match	100.0%;	Score 1729;	DB 9; Length 1729;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1729;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

Db	1	CCCTGCTCTCCCCCGAGTCCCTTTTTCCTTCTCTCTCTTATTCCTT	60
Qy	61	tcaataacatattcttctctctctaggtttcaattcaacttgaaatagtggttaa	120
Db	61	TCAATAACAAATATCTTCTCTCTTAGGGTTTCACTTCACTGAAATAGTGTATA	120
Qy	121	atgaatgatatgttaggtctcttggtcttcaataatcaatcaactcaatcatga	180
Db	121	ATTGAATGATATGTAGATTCCTTTGGTTTCTCATCATCTCACTCACTCACTCATGA	180
Qy	181	tgaagaattcttcgatacatcaatacaacgylagatltcaacgctactgtcaaatcaac	240
Db	181	TGAAGAATCTTTCGATCATCATCAACAGTAGATCAACGCTACTGCTCAACTATCAC	240
Qy	241	cacacacacatctactactcaacccgctatcaaatccgcagatctaatctagctgtac	300
Db	241	CACACACACATCTACTACTCAACCAAGCTATTCGAATCCGCAGTACTCATTAAGCTGTAC	300
Qy	301	ggagtgatttcgactgcgaattctcgcgcgcgcgaagaagactccctactatattatcaac	360
Db	301	GGAGTGAATTTCCGATGCCGATTTCTCGCGCGCGAAGAAGACTCCTTACTATATTATCAAC	360
Qy	361	taactcatctcccttttgggtatcaactgaacggttagtcataatattactcgccact	420
Db	361	TAACTCATCTCCCTTTTGGGTATCAACTGAACGGTTAGTTCATTAATTTCTGCGCACT	420
Qy	421	ttcccttcgctcaacgcgcatatatacgcgcaaccccaatattctagaacctgtlga	480
Db	421	TTCCCTTCGCTCAACCGCATATATATGCTCAACCAATATTTCAATGACACCTGTGA	480
Qy	481	aacaactcaactgatatctctctcttcgcatcaattagctcattatcaatcatatct	540
Db	481	AACAACCTCAACTATCTCTCTCTTCGCTCATCATTAAGCTCAATTAATCATCATATCT	540
Qy	541	atctctaaacaaattaccccttctagaatttcaatcaatcaacgctaaacgaat	600
Db	541	ATCTCTAAACAAATTAACCTTTCTCAATAGGTTTACTCATTAACCGGTATCAAGGAT	600
Qy	601	tttagaagcgatlttaacgylaatcaatcaagaacatccaatcgltgatttcgacattca	660
Db	601	TTTAGAAGCGATTTAAGGTATATCATCAACCAATCAATCGTTGATTTCGACATTAATCA	660
Qy	661	cggggttcaatggccaccgttaatgaagcactagatgactgaacctgtcccactc	720
Db	661	CGGGTTCAATGGCCACC GTTAATGAGAAGACTAGCTGATCGTTACCTGCTCCACTCT	720
Qy	721	tcgaaatccggtactggaatgaacttgaacctgtlagaacaggtlgttagc	780
Db	721	TGGAATCCACGGTACTGGAATGAAATGACCTGTATACCTTGTAGAACAGGTGATCGTTAGC	780
Qy	781	taattgtcactcaatcattaggttagatttcaatctcaatctctcttataagcaata	840
Db	781	TAAATTTGCTCACTATTAAGGTGGAATTTCAATTCACCTCTTTTATATAGCCAAATA	840
Qy	841	taacaacatcaagatgaagaatcttatattcttccatctgtactcaactccctgtga	900
Db	841	TAAACACGATCAAGATGAAGATCTTATATTCTTCATATTGTAACATACCTCGATGA	900
Qy	901	aacctagctatcaactgttcttctactcccaacgcgcctttaaagaacgcgaaagt	960
Db	901	AACCTAGCTATCAACTGTGTTTCTCACTCCACCGCTTTTAAAGAACCGGAAAGTT	960
Qy	961	aagatatttttgcataaggttaagtaataagaacctaaatattgttaaatcgagaa	1020
Db	961	AAGATATTTTTCATAGGTTAAGTAAAGAACCCTAAATATGTATCAATCGCGAGAA	1020
Qy	1021	ggaagaataataaccaatccctcttttatacaagaattcaatcagagcggttgattalia	1080
Db	1021	GGAGAATAATCAATCAATCCTCTTTTTCACAAAGATTCAATGAGCGCTTGATTATTA	1080
Qy	1081	taagctgttttgatttcaactggaagctaacatgccaacgcggttagtcgaagagatgac	1140





```

introns
complement(10449..11057)
/number=3
exons
complement(11058..11266)
/number=2
introns
complement(11267..11865)
/number=1
complement(11866..12087)
/number=1
exons
25598..29176
25598..29176
/number=1
/notes="ORF2"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC36403.1"
/db_xref="GI:13620228"
/translation="MSHRCHEPFLSLGVKKNVEMILLSSFLVYVSNARKLAEND
OVRLLAFKSSVSDNGFLNFWLTLSSSPCTWNGISCSGQVVELNSVGLSGL
HLDMLPFLRLVNFSGNHYGLSLSSASCSPEFLDLNANFSEVLEPLKSGD
NIKLYANSKISGVKLPKPSILQDLSSNFIISDPLISALSNCONLNLNFSNK
IACKLSISCKSLSVLDSRNLTGELNDIDGTCONLTVLNSFNLTVEPPS
LANQSLNLTIANISIRMETPVELLVKLSLKLVLAHNQFDKIPSELGQSCSTLE
ELDSGRNLGELPSTKLSLSEFLMGNNEISGDELNTVSSLTMLRLYLPNNI
TGVPKSLVCTKQLVQLSMAFIGNVSEPCASGFPLETMLASNYLGVPPK
LGHCRLRKIDLSFNMLVGSIPLEIMLPLNSELVMANNLTGEIPGICINGNLQ
LIANNFISGTLPOSISCTNLMVSVLSLSNGEIPGICINLANLAILOQNSITG
PIRGLSGSRLLHDLNSNALTSTISLELADQGHVPMASGKORAFVNEGTEC
RGAGLVEFEGIREERLALIPMHFCPSTRISGRTMYTFTSNGSMIYLDLSYSLG
TIPNDLSLFLQVNLGNHNFCTGIPFNEGLKIVGLDLSHNSLOGFTPPSGLS
FLSDLVSNMNLSTIPSGQLTPSPASRYENSLGCVLPFGSGNGHSSSIYH
GNKPTTIGMVGIVMSEFICITLILVALYKIKTONEBERKDKIDSLPTGSSSMK
STVEPESINAVTEFKPLRKTLEHATNGFSSSMIGSGFGVYKQDLRGSVY
AIKTYHTGQDREPMEMETICKIKHNVPLIGCYCKISEBLLIYEYKMSLS
VLHGGGCMFLDMPARKIATGSRGLATLHSCIPHILHRDKSSVGLDENEAR
VSDGKGLVNLADLTLVSTLACTPGVPEYQSECTAKGQVSYVGLLELSC
KRPIDPVTMGDDNMLVGMKQHLNDKOSHELIDELITNLSGDAELHYLKVARECD
EKSKRPMTIOVMTKFEVQDSESDILDISVKSILSESQEREP"
39530..39982
/notes="ORF3"
/number=1
join(39530..39982,41510..41698,41775..42637,42740..42898,
42994..43549)
/notes="ORF3"
join(39530..39982,41510..41698,41775..42637,42740..42898,
42994..43549)
/notes="ORF3"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC36402.1"
/db_xref="GI:13620227"
/translation="MGGFDDHVAIMGDMMPSPSPRFFSSLLDDGGSRLSPGCTNE
TSGMLASGPOHVGNSDGVDTAPAIASEPOVKLSVPSDOMKNSRGGLERIARAG
FNAPKLTESLIRPADLRNOGVRSPYLTIPGLSPFLSPLSVLPSPPTK
FLFSSGIDNRSLAMEDTDKRNDALESISSEFSPKPYEETAPSLFPTGSSMTQ
VNPNFSQOQPNIEVSHSONSLVSHRVETATONGTLOQSSDPREFGAEDVMA
NNVTLDSRTQGVSAVDHSPRLDEPDEDIDOGGDPANVAGAPADGVMMRYGOK
QYKSEYPRSTYKTHPTCYKPKKVSPEHTEITIKKHNPKPPNRKSLGST
NLSLDLVYDGAEGASVNGDLGOANFHKAAGGGGEGEVRNNNDANLGSCHRASAP
FSAQNTLRLEDAVDVSNFSDNEDEDRGTHSVQGEDEDEDSKRRKLETYS
ADMTGATRAIREPVVQTTSEVDILDDGYWMRYGQKVVGNPNPSSYCTSGACN
VRKHVERASHDLKVVITTYEGKHNHVDPAANSHTVNSVAGNPASTVTAQONLHR
PEAOLONAMARPDROPSLGSFGLSGRPDGPMPGSGYNOGOGSLSLAMAFPHN
QNRGEVPMHRYLQPRPMHMGPMFKPEKXVPLSDPGTATAOFLNLSLGPOM"
39983..41508
/notes="ORF3"
/number=1
41510..41698
/notes="ORF3"
/number=2
41699..41774
/notes="ORF3"
/number=2
41775..42637
exons

```

```

introns
/notes="ORF3"
/number=3
42638..42739
/notes="ORF3"
/number=3
42740..42898
/notes="ORF3"
/number=4
42899..42993
/notes="ORF3"
/number=4
42994..43547
/notes="ORF3"
/number=5
complement(44881..44962)
/notes="ORF4"
/number=9
complement(join(44881..44962,45750..45982,46133..46345,
46506..46736,48694..48798,48892..49346,50526..50720,
51517..51919,52622..52957))
/notes="ORF4"
complement(join(44881..44962,45750..45982,46133..46345,
46506..46736,48694..48798,48892..49346,50526..50720,
51517..51919,52622..52957))
/notes="ORF4"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC36403.1"
/db_xref="GI:13620228"
/translation="MSGGEYDHNILRLSTSSASEGVESQSPRRNTSRISKDLNR
LDREFSGSRDRHNNHSSSPSRNRGVSTADEILDSDAPPMAMLVGCLGLAT
LDVAFGRNGVHIVHEMAMACTPYDGAAMLRLOLADPMHRIILLIPVIGVYGMNG
LIGLIDLTQSSSTOGGCFDIACITPVYKTOAIIITGCSLGPBEPSPDIKSCA
TGCMMENNERNRIALIAAGAAAGTSGFNAAGSFFALEYLRRLRAVNSPFTY
AMIIASVISTVNAVNGEKQAFVPTPYDKSAEPLYLILMLGCVSVETRLV
SWFKGQFLKEKGLSDVYCPALGGLAGVIALRPIGLWGTNDELITGKTAS
AGIGMELQALVAALVAVATLCKSGVLGLAPSLMIGAAGVAGGLAGELISAIP
GNAAIAOQAAIVAGMAATLASVCSPTSVLLPELTKDRIIILPLMGVYGLAIWV
SYTPQPNDAESSEAKFASKYSITSPDENNEGSRSGERNLLEKVNHSHESD
EGLILEDLKYSQANSNDLKYSPQYKALECHMDGROSCLVYVDEGLYEGILITIG
DVKRSLEFNHGDSSNKLSDVNDANTCVSSICTGKISYRGDCGLTCTYPTDLAIK
QLMKAKGILKQLPVYVRGGEFRRERRRRIALLHYDSVEETIRREVSHRSKYQONEE
KDKOITNGH"
complement(44963..45749)
/notes="ORF4"
/number=8
complement(45750..45982)
/notes="ORF4"
/number=8
complement(45983..46132)
/notes="ORF4"
/number=7
Query Match 100.0%; Score 1729; DB 14; Length 57047;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

Db	2399	TTATCCTTCTGAAGGCTATCACTCGGAGTTTCGAGTAATCTTCTCTTAGGTGCA	2940
QY	1381	aaatcaacccccctttctccatctcgtctgtgcgltgagaaaaacataatagccaact	1440
Db	2939	AAATCAACACCCCTTTCTCCATCTCGTCTTGCGGGTGAGAAAAACATATAAATAGCAACT	2880
QY	1441	tcaaagaggtaataaagacatactagtttgaagagatctgaagaaacgcgttgagtg	1500
Db	2879	TCAGAGGATATTBAGACTACTGATAGTTTAGAGGAGATCTCAAGAAAAACGGGTGAGTG	2820
QY	1501	aaacccctaataacacagattttctaaatgaagtgtgtagtagaagaatttcgacgtgaa	1560
Db	2819	AAACCCCTTAATTAACCGATTTTCTAATGAAGTTGTAGTAGAGAAATTTGCATGCTGAA	2760
QY	1561	gaacacatattgaagagatgatgaattcaatgctttcttgattactatgatgaatgaa	1620
Db	2759	GAAACAATATTCAGAGGATATGAAATTTCAATGTTTTTTTGTTTACTATATGATATGAA	2700
QY	1621	tgttttaaaatttttaacacagagagacagtgtgatgatataatgatttaagtaacta	1680
Db	2659	TGTTTTAAATTTTTTAACATAGAGAGACTAGCTGATGATATATAGTAATTTAAGTTAACTA	2640
QY	1681	gtctttgtataacgacagatcttgatataactatttttattttaata 1729	
Db	2639	GTCTTTGATATACGCAAGACTCTGATCAACTATTTTTATTTTAATTA 2591	

	RESULT	4		PAT	21-JAN-2000
Dd	LOCUS	A84080	1296 bp	DNA	
	DEFINITION	Sequence 9 from Patent WO9846759.			
	ACCESSION	A84080			
	VERSION	A84080.1	GI:6733220		
	KEYWORDS				
SOURCE	potato.				
ORGANISM	<i>Solanum tuberosum</i>				
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum. 1 (bases 1 to 1296)				
REFERENCE	PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED ABSCISSION AREA FORMATION Patent: WO 9846759-A 9 22-OCT-1998;				
AUTHORS	Theres, N.				
TITLE	THERES NILOLAUS (DE) Location/Qualifiers 1..1296 /organism="Solanum tuberosum" /db_xref="taxon:4113"				
JOURNAL					
FEATURES	source				
BASE COUNT	354	a	303	c	229 g 410 t
ORIGIN					
	Query Match	70.7%;	Score 1223.2;	DB 9;	Length 1296;
	Best Local Similarity	97.1%;	Pred. No. 3.2e-251;		
Matches 1259;	Conservative	0;	Mismatches 28;	Indels	9; Gaps 1
QY	131 atgttagatcccttggtttcataatatcctaaccctacatgatgaagaacct	190			
Dd	1 ATGTTAGAACCCTTTGGATTCTTCATCATCAATCACAATCACCCCTATGTAAGAAGACT	60			
QY	191 tcatacatcatacaacgcgtagattcacgccactgctactgtaacaactaacaccaccacc	250			
Dd	61 TCTGATCATCATCAACGCCGTAGATTACCGCTACCTCFACAACTATACACACACACACC	120			
QY	251 atcatacctoacagctattccaatcccgccagtactatttgctgltggagggttgatt	310			
Dd	121 ACAACGACACCTCACGCATATTCAAATCCGCCACCTACTCAATTAGCTGTGGAGATTGATT	180			
QY	311 tcgaagtcgagattcttgcgcgcgaaagaagacctctcatatatataaatcaatcatct	370			
Dd	181 TTGGGGTCCGATTTCCTGGCCGCCAAAAGACTCCTTGCAATTATATCAACTAACCTGTCT	240			

QY 371 cctttgtgatacaactgaacgggttaagtcacattactcgcgcacttccttcgt 430  
 |||||||  
 Db 241 CCTTTGGTGTGATCAACGGAACGGTACTCCATCAGTTTACTCGCGACTTCCCTTGT 300  
 QY 431 ctcaacgcgtatataatcgtaacacccaatcatttcagacacctgttgaacaactca 490  
 |||||||  
 Db 301 CTCACACCGCTATATATGCTCAACCAACCATTCATTCATGACACCGTGTGAACAACCTCA 360  
 QY 491 acgatctctctct-----tcgtacatattagctcctaattcaatcatalctta 541  
 |||||||  
 Db 361 ACTGATCTTATCTTCTGCTCCATCGCATCATGATTAAGCTCTTAATTCATCATCATATCAT 420  
 QY 542 tctctaaacgaattacacctcttcataagtttactcaatlaaacgcgcaacgaacgatt 601  
 |||||||  
 Db 421 TCTCTAAATCAAGTTACCCCTTTTATAGGTTTACTCATTTAACCGCTTAATCAACGCAATT 480  
 QY 602 ttagaagcgatlaacggtatcatcaacgaacaaatcaacatcgattgatttcgacatlaacac 661  
 |||||||  
 Db 481 TTAGAAGCGATTAACGGTATCATCAACCAACCATTCGTTGATTTGACATTTATATCAC 540  
 QY 662 ggggttcaattggccacggttaatgcacgactagctgattcctcgtcccaactct 721  
 |||||||  
 Db 541 GGGGTTCATGCGCACCGCTTATGCAACACTAGCTGATCGTTACCTGCTCCATCTCT 600  
 QY 722 cgaatcacggttactggaatgacacctgataacctcgttagaagcggtgacgtttagct 781  
 |||||||  
 Db 601 CGAATCACCGGTACTGGAATGACCTTGTATACCTCTGATACAGAGTGATCGTTTACT 660  
 QY 782 aaattgctacatcaattaggggttagattcaatcatalctccttatalatagccaataa 841  
 |||||||  
 Db 661 AAATTTGCTCACTATTTGGGTTGAGATTTCAATTCATCTCTTATATGCGCAATTAAT 720  
 QY 842 aacacagatcagatgaagatccctcttatttctcctcattgacttactcctcgatga 901  
 |||||||  
 Db 721 AACCGCGATCAGGGAAGATCTTATATTTCTCCATTTGATCTCTCCCTGATGAA 780  
 QY 902 accctagatcaactggtttcttctacccacgcgacctttaaagaacgcggaagaatla 961  
 |||||||  
 Db 781 ACCCTAGATCAACTGCTTTTCATCTCCACCGCTTTTAAAGACCGCAAAATTA 840  
 QY 962 aggaatttttgcataaggggttaagtaaacctaaatgttacaatcgcggaagaag 1021  
 |||||||  
 Db 841 AGGATTTTTCATAGGAGTTAAGCATGAACCTTAATTTGTTACAAATGCGGAGAAAG 900  
 QY 1022 gaagcaaatcataacacacctcttctttaaagaattcagagcggttgattatata 1081  
 |||||||  
 Db 901 GAAGCAAAATCAACCATCTCTTTTACAAATTTATCGAGGCGTTGATTTAT 960  
 QY 1082 aacagctggtttgattcactggaagctacatctgcaacgcggttagtcgaagagatga 1141  
 |||||||  
 Db 961 ACAGCTGTGTTGATTCATTTGGAACCTCATCTGCAACCGGATGCTGAGAGATGACA 1020  
 QY 1142 gttagaacaagtgtggtttggaagagagatgtgatalcgttgcgattggaagaagataa 1201  
 |||||||  
 Db 1021 GTTGAACAAGTGTGTTGGGAGAGAAATTTGTTGATTCGTCGATGGAAGAGATATA 1080  
 QY 1202 aggaagaagaacatgtaaaagtttagatcatcagtggaagtatgttgagagattggaatt 1261  
 |||||||  
 Db 1081 AGGAAGAAGAAACATGAAAGGTTTATGATCATCGGAAGTATGTCAGAGATGTCGATT 1140  
 QY 1262 agtaagtgtggtttaaagcccttgcattatcaacgaagtaagctctttagaactat 1321  
 |||||||  
 Db 1141 AGTATATGTTGCTTTAAGCCCTTTTGCAATATACACAGCTTAAGCTTTTGAAGACTACAT 1200  
 QY 1322 tatccttcgaaggtatcaacatcgagatttcgagtaattcttcttccttagtggtaa 1381  
 |||||||  
 Db 1201 TATCTCTGTAAGGCTATCACTCGAGATTCGAGTATTTCTTCTTCTTGTGGCAA 1260  
 QY 1382 aataaccccccttctcattcattcgtcttgcggttga 1417  
 |||||||  
 Db 1261 AATCAACCTCTTTCTCATCTCGTCTGGGCTTGA 1296

RESULT 5  
 AC005223/C  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana chromosome I BAC T5A14 genomic sequence,  
 complete sequence.  
 AC005223  
 VERSION  
 AC005223.1 GI:4096078  
 HG.  
 SOURCE  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE  
 AUTHORS  
 1 (bases 1 to 87967)  
 Federespiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
 Gonzalez,A., Kremetskaya,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
 Luros,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaya,V.S.,  
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
 Unpublished  
 2 (bases 1 to 87967)  
 Federespiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,  
 Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,  
 Oji,O., Osborne,B.I., Shinn,P., Sun,H., Tortum,M., Vyotskaya,V.,  
 Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (01-JUL-1998) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 3 (bases 1 to 87967)  
 Federespiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
 Gonzalez,A., Kremetskaya,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
 Luros,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaya,V.,  
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (05-JAN-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 4 (bases 1 to 87967)  
 Federespiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
 Gonzalez,A., Kremetskaya,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
 Luros,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaya,V.,  
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (30-JAN-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 COMMENT  
 On Jan 5, 1999 this sequence version replaced gi:3779015.  
 Bases 1-9066 of clone T5A14 overlap with bases  
 100,099-109,160 of 'IGF' clone F20N2, gb/AC002328.  
 e-mail for correspondence: arabes@sequence.stanford.edu  
 Genes with similarity to proteins in the databases are described as  
 'putative', '-like' or 'similar to'. Genes that have EST  
 similarity but no significant protein similarity are described as  
 'unknown proteins'. Genes that are annotated based only on gene  
 prediction software are described as 'hypothetical proteins'.  
 The software programs used to predict genes include: Graal  
 (Informatics Group, Oak Ridge National Laboratory,  
 http://complib.ornl.gov/section/index.html), GENSCAN (Chris Burge,  
 http://genomic.stanford.edu/~chris/GENSCAN.html), Fexa (V.S. Solov'yev  
 & A. Salamov, Sanger Centre, http://genomic.sanger.ac.uk/), and  
 NetPlantene (S.M. Hedgesgard, et al., CBS, Technical University of  
 Denmark, http://www.cbs.dtu.dk/NetPlantene.html).  
 FEATURES  
 SOURCE  
 1. 87967  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /chromosome="I"  
 misc\_feature  
 1. 9066  
 /note="overlap with bases 100,099-109,160 of 'IGF' clone





Oy	402	atcaattactcgcaacattcccttccttcgtctcaacggcta-----	441
Db	6583	ACCTCTTCACATAAAGCCTTCTCCGTCAGATACACCGTGTGACACACAGATCAGACGGCTG	6534
Oy	442	-----tatatgctcaaccaccaatcattatgatgacacgtgtgaaacaac-----	486
Db	6523	AAACGTTGGCCAGTGGACAGCAGACGAATGACGATGAGTAATCCAGCGGTTCCAGCA	6464
Oy	487	-----tccaactgattcttcttcctgcgtcatcatatgctctaatcatc---	534
Db	6463	GCAGTGTATCCAAAGAACAGTCTTGTCTTTCGAAACCAAGAACAAATCTGCATTCGAGT	6404
Oy	535	-----atactatcctcaaccagaattaccocctctcaaaagttaaggttactcaatlaacg	587
Db	6403	CTTGTTACTATATCTTTGGCTTAACCACTAACGCCGCTTATTTGGTTGGTGTATTAACGG	6344
Oy	588	ctatacaagcgattttagaagaagattaacggtatcatca---agcaatccacatcgttg	644
Db	6343	CGAACCAAGGTATCTCTCGACGCGACGAGAACAAACGTTAACGACCTCTACATTATATCTG	6284
Oy	645	attcgcgacattaacaggggttcaatggtgcacoggttaagtgaacgacttagtgcgtt	704
Db	6283	ATTTAGTATATATCACAAGACCTTCATTTGGCTTCATTTATGCAAGCCTTAGCAGAGAGT	6224
Oy	705	ac-----cctgtcccaactcttcgaatcaacggtacgtcgtgaatgcgttg	749
Db	6223	CATCAAAACCTTAGACGTACCTCCATCTCTCCGCTACCAAGGATCGGTCGACATGTAA	6164
Oy	750	ataccctctgtagaacaagcgtgacgttgcttaaatcttctaccattcattagtggtgat	809
Db	6163	CCGGATTAAACCGAATCGGAGACCGGTAAACCCGGTTCCTACATCTTTAGATCTCCAAAT	6104
Oy	810	tccaatccatccctcttataatagccaataataacaacgatacagatgaatccctcta	869
Db	6103	TCCAGTTTCAACGCGTATGTATGTGTAAGAAGATCTCCCGGACTTTTGCTACAGATCC	6044
Oy	870	ttattccctcatgtactacccctgatagaaaccttagcatcaactgltgtttctaac	929
Db	6043	GATGTAGTCTCTCAGCCGTCACAAAGGAGACCTTCGCTCAATGTGTTCACTTCC	5984
Oy	930	tccacgcgcctttaaagaac---cgagaaagtaagattttttgcaiaagggttaagt	986
Db	5983	TCACCAAAATATTTTACGACGATGAGATATGTATCGGTACTCTTGTGCAGGCATCAAGA	5924
Oy	987	caatgaacccttaaatatgtttacaatcgcgcgagaggaagacaatataaccaatccctctt	1046
Db	5923	GCTTAACCTCTAGATATGTTTACATATGGCAGAGAGACAACTAATCATGTGAGATCACTGT	5864
Oy	1047	tttaacaagaattcatcgtgagcggttggtattatatacaacgtgtgtttgattcaatcgtga	1106
Db	5863	TCTTGAATAGATTTCTCGAGGCAAGTGGATCATTAACATGCGATCTTTGATTTGTTGGAG	5804
Oy	1107	ctacatttgcacccgcggtgaagcgcgagagagatgaacattgaacaagcgtggtttggaag	1166
Db	5803	CGAGCTTGCCGCCAAATATAGCCGAGAGAGACTAACCTTAAGCAACGGTGTCCGGTAAAG	5744
Oy	1167	agatttgtatatacgtgtgcgaatggaagagataaagaagaanaagacaatgaaggttta	1226
Db	5743	AGATTTTGGATGTGTGCGCGCGAGAGACGAGAGGAAGCAAGCAACATCTCGAGAGTTTG	5684
Oy	1227	gatactgggaagatgtatgaaggaattgtggaattgaagtgttgccttaagccctttg	1286
Db	5683	AGATTTGGGAGAGATATTAAGAGTTTGTTTCTTTAACGTTCTCTATTGGAAACCTTGG	5624
Oy	1287	caatacaacaagctaagcctcttlttgagactcatatcctcttgaaggctatacaactcg	1346
Db	5623	CTTTGTCTCAAGCTAACCTCTCTTACACTTATATCTTCAGAAAGGTATATATCTTC	5564
Oy	1347	gaagttcaggaatcttcttctcttggttggtgaataataaaccccttctccatcctgt	1406
Db	5563	AGTTCCTTAACAATCTTGTGTTCTTGGGCGGGAATAATGTCCCTCTCTCCGTTTGTG	5504

oy	1407	cttggccttggagaanaa	1423	
db	5503	CGTGGAAATGATCCAA	5487	
RESULT	6			
LOCUS	AC002328			
DEFINITION	Genomic sequence for <i>Arabidopsis thaliana</i> BAC F20N2 from chromosome			
ACCESSION	AC002328			
VERSION	AC002328.2			
KEYWORDS	HTG:			
SOURCE	thale cress.			
ORGANISM	<i>Arabidopsis thaliana</i>			
REFERENCE	Embryophyta: Viridiplantae; Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
AUTHORS	1 (bases 1 to 109171)			
REFERENCE	Khan, S., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Kim, C., Shinn, P., Altafi, H., Bel, Q., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howng, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharshy, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thavel, A., Toriumi, M., Vaysberg, M., Yu, G., Federspiel, N.A., Theologis, A. and Ecker, J.R.			
AUTHORS	1			
TITLE	Genomic sequence for <i>Arabidopsis thaliana</i> BAC F20N2 from chromosome			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 109171)			
AUTHORS	Ecker, J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-JUL-1997) <i>Arabidopsis thaliana</i> Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA			
REFERENCE	3 (bases 1 to 109171)			
AUTHORS	Ecker, J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-AUG-1998) <i>Arabidopsis thaliana</i> Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA			
REFERENCE	4 (bases 1 to 109171)			
AUTHORS	Shinn, P., Buehler, E.E., Dunn, P.P., Feng, J.J., Kim, C.C., Li, Y.Y., Walker, M.M., Altafi, H.H., Araujo, R.R., Conn, L.L., Conway, A.A.B., Gonzalez, A.A., Hansen, N.N.F., Huitzer, L.L., Kremenetskaia, I.I., Lenz, C.C., Li, J.J., Liu, S.S., Lucos, S.S., Rowley, D.D., Schwartz, J.J., Toriumi, M.M., Vysotskaia, V.V., Yu, G.G., Davis, R.R.W., Federspiel, N.N.A., Theologis, A.A. and Ecker, J.J.R.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (03-DEC-1998) <i>Arabidopsis thaliana</i> Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA			
AUTHORS	5 (bases 1 to 109171)			
REFERENCE	Ecker, J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-MAY-2000) <i>Arabidopsis thaliana</i> Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA			
REFERENCE	6 (bases 1 to 109171)			
AUTHORS	Chen, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bel, B., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howng, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharshy, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thavel, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.			
JOURNAL	Direct Submission			
COMMENT	Submitted (28-JUN-2000) <i>Arabidopsis thaliana</i> Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA			
FEATURES	On May 13, 2000 this sequence version replaced gi:3492855.			
SOURCE	Location/Qualifiers			
	1..109171			



Oy	342	tccttactatattataaactaacatctctcccttttggatctcaactgaaggttagtgc	401
Db	2424	TACTTCATACTCTCTCCCTTAACCTCTTCTCTCAGCGGACCTCCAGAGACTTGAC	2483
Oy	402	atcaatttactcgcgcactttccctctcgtctcaaccgcta-----	441
Db	2484	ACCTCTCACTAAAGCCTTGCCGTACGAATCAACGCGTACGAACAAGATCAAGACGGTG	2543
Oy	442	-----tatatgctaacccaacatcatttcaagacacgtttgaacaaac-----	486
Db	2544	AAACGTTGGCCAGCTGTGACGACCAAGAAATGACATGAGTAACTCACGGGTTCACGA	2603
Oy	487	-----tccaaactgattctctctcttcgcatactaatgagctctaatcaatc---	534
Db	2604	GCACTGTATTCGAAGAAGCAAGTTCTTTTCGAAACCAAGACACAACTTCTGACTTCAGAT	2663
Oy	535	-----atatctactctaaaccaagttaccctcttcaataagttactcaathaacg	587
Db	2664	CTTGTACTTCTTTGGCTTAACCAACTAACCCCGTTTATTTGGTTCGGTCAATTAAACGG	2723
Oy	588	ctaatcgaagattttagaagcgatataacgtaacalca---agcaatccaactcgttg	644
Db	2724	CGAACCAAGTATCTTCGACGCGACGGAACAACAGTAAACGACTCTCAATATACTTG	2783
Oy	645	atttcgacatlaatacggggttcaatgtgcacacgttlaatgaacactagctgactt	704
Db	2784	ATTTAGATATATCCAAAGACCTTCATATGGCTCCTCATTTATGCAACCCCTACGACAGAGGT	2843
Oy	705	ac-----cctgtctccactctcttcgaatcaacggtactcgtgaattgactcttg	749
Db	2844	CATCAAAACCTTAGACAGTCCACCTCCCTCTCTCCGATTAACCGGATCGGTCGAGATGTAA	2903
Oy	750	ataccctctgtagaacaggtgactcgttttagcttaaatcttctactcatatgaagttagat	809
Db	2904	CCGATTAAACCGAATCGAGAGACCGGTTAACCGGTTGCGTCACTCTTAAAGTCTCCAAAT	2963
Oy	810	ttcaatctcatctcctttatatagccaataataaacagatcaagatgaatcccttcta	869
Db	2964	TCCAGTTTCAACGCTAGTATGATGTAAGAAAGATCTCCGCGACTTTTGGTACAGATCC	3023
Oy	870	ttatttcccatattgtaactactcccgtatgaaaccttagctataactggtttttctacc	929
Db	3024	GATTGTTAGCTCTCTCAGCCGCTACAAAGAGACACCTTGCCGTCAATGTGTCTACTCTCC	3083
Oy	930	tccaacgcctcttlaanaagac---cgcgaaaagtllaaggaatttttttgatagggttaa	986
Db	3084	TCCACAAAATATTTTAAGCAGACGATGAGATATGATGCTACTTCTTGACGAGATCAACGA	3143
Oy	987	caatgaacccctlaaatttgttaacaatcgcgggagaaggaagaatcataccatcctcttt	1046
Db	3144	GCTTAAACTCTAGAACTGTTCATCATGAGAGAGAAACCTATCATGAGATCACTCTGT	3203
Oy	1047	ttttcaagaatcatcatcgagcgcttgatattatatacaagctgtgtttgatccaaggaag	1106
Db	3204	TCTTGATTAAGTTCTCTCGAGCAGCTGGATCATTAATGGCGAATCTTTGATTCGTTGGAG	3263
Oy	1107	ctacatttgcacacggtagtctgagagagatgaacagttgaacaagtgigttyggagag	1166
Db	3264	CGAGCTTGGCCGCAAAATAGCCCGAGACAGACTTAACCTTAAGCAACAGGAGTTCGGTAGG	3323
Oy	1167	agatctgttgaatcgtttgcaatggaagggatataaaggaaagaagaatgaaggttta	1226
Db	3324	AGATTTTGGATGTTGTGGCGGCGGAAGACAGGAGAGAAAGCAAAAGACATCGAGAGTTTG	3383
Oy	1227	gatcaatggaagtatgattgagagagttgtygaatttgaatattgttcttlaagccctttt	1286
Db	3384	AGATTTTGGGAAGAGATGATGAAGAAGTTGGTTGCTTAAAGCTTCTTATGGAACCTTGG	3443
Oy	1287	catatacaagaactaagctctcttttgaagcttcaatattatcctcttgaaggtcatacaacg	1346
Db	3444	CTTTGTCTCAAGAGCTTCTCTTAAAGCTTCATATATCTTCAACAGGATATATATCTTC	3503

QY	1347	gagttgcgaagaaatcttcttccttaggttttgcaaaatcaacccctttccatccgt	1406
Db	3504	ACTGTCCTTAACAAATCTTTGTTCTTGGCTGGCAAAATCGTCCCTCTTCCGTTTCGT	3563
QY	1407	cttgagcttgagaaaaa	1423
Db	3564	CCTGGAATGATCCAAA	3580
RESULT	7		
LOCUS	CRU303349		
DEFINITION	Capsella rubella ORF1, ORF2, ORF3, ORF4, ORF5 and ORF6 (partial).		
ACCESSION	AJ303349		
VERSION	AJ303349.1		
KEYWORDS	ORF1; ORF2; ORF3; ORF4; ORF5; ORF6.		
SOURCE	Capsella rubella.		
ORGANISM	Capsella rubella		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Capsella.		
AUTHORS	Rosenberg, M., Theres, K., Acarkan, A., Herrero, R., Schmitt, T., Schumacher, K., Schmitt, G. and Schmidt, R. H.		
TITLE	Comparative sequence analysis reveals extensive microcolinearity in the lateral suppressor regions of the tomato, arabidopsis, and Capsella genomes		
JOURNAL	Plant Cell 13 (4), 979-988 (2001)		
MEDLINE	21178822		
REFERENCE	2 (bases 1 to 27056)		
AUTHORS	Schmidt, R. H.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JAN-2001) Schmidt R. H., Max-Delbrueck-Laboratorium in der Max-Planck-Gesellschaft, Max-Delbrueck-Laboratorium in der Max-Planck-Gesellschaft, Carl-von-Linne-Weg 10, 50829 Koeln, GERMANY		
FEATURES			
source	Location/Qualifiers		
	1..27056		
	/organism="Capsella rubella"		
	/db_xref="taxon:81985"		
	/clone_lib="cosmid library"		
	1616..2959		
	/note="ORF1"		
	/number=1		
	1616..2959		
	/note="ORF1"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAC36387.1"		
	/translation="MTLSFKSSSSSEDEATETPPPCQLASSSSSSAATSASHLRRLFTPADEFVSOMFTARNLSTLSLNSSPYGSSTORLAHLPKRLSRNRLOOORODPYATCTDNTMTSTNTSTVTSSVCKBOPLFRKNNNSDEQCYLMTLNDLPFIRFCGLRANQALIDATETENDGALHIHIDLDISGLQPPPLMQALAESSSPSPSRITTGCGDVGVLNRTGRLTRFANSGLQFRLVLVDEDEPLLILRLALSAVGELTIAVNCVLYLFNFDDADLDLGHFLTVIKSLNPRTVAERANGDHSFLIRSEALDHYVAIFSLSEATLPPNSRERLTLEQPMFGKIMDVVAATERKORRRREIFGEMKKRREGVNVNIGSFALSQAKLRLRHYPSGVNLFQPLNDSLFLGWNRLLEFSVSKR"		
	1616..2959		
	/note="ORF1"		
	/complement(join(7709..9082,9175..9365,9593..9798,9871..10020))		
	/note="ORF2"		
	/complement(join(7709..9082,9175..9265,9593..9798,9871..10020))		
	/note="ORF2"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAC36388.1"		
	/db_xref="GI:13620167"		
	/translation="MEETRSVLAEEILKRLDLENCVACVSTLRSAVSGVLSLTSLDVSFSPDEETLNHYRGCIGLSITLNLCELDLANSRVGVLGPHLREHLRGLTSSVLSIGTLCENLRVLTLEMSDLDSPAFQSNLTOMLNGCYLESIDINIRGILV		
CDS			
exon			
CDS			
mrna			
mrna			
CDS			

exon	complement(7709. .9082) /note="ORF2"	DATAFQSVRESLPETLTKALRLOPLESEAILMNRKYTGCTLSQADVSLAISSPSD TASISLSTVILDIISDRLLIITAITGSLPOLVOLJEDREPREKEPPFDSDLTITGQALAGTCO QULSLSIVRCYVNRKISFRINDMGJFLLISEACKGLESVRLGGFOKVSADGASLASH CRNKKRKEVSGCALISDLAEFDVYTGSSCFLOEYRISTCPLITSEAVKRIKGLGNCLEVL DIGCKRISIPSCINSVYLRRLSLINAGADVDSMLAIGKSDVPIIQLSIRGKRLV SDRISLHLLINBETISIKLSTLIDLGHPMGMSDRAHTITIRCKALTELSIRCFVTD SSTLSLAEVTEGQVEGGSKOLRKANINCVSLTGGALRMLSKRSPAGLHMLGLGQTRIA GRKEVYATTCGQGPWLLTLCDFCGELGCTDGMFEHFIPOKH"
intron	/number=4 complement(9083. .9174) /note="ORF2"	
exon	/number=3 complement(9175. .9265) /note="ORF2"	
intron	/number=3 complement(9266. .9592) /note="ORF2"	
exon	/number=2 complement(9593. .9798) /note="ORF2"	
intron	/number=2 complement(9799. .9870) /note="ORF2"	
exon	/number=1 complement(9871. .10020) /note="ORF2"	
exon	/number=1 complement(10491. .10953) /note="ORF3"	
mRNA	/number=5 complement(join(10491. .10953,11109. .11264,11359. .11761, 11859. .12000,12520. .12897)) /note="ORF3"	
CDS	complement(join(10491. .10953,11109. .11264,11359. .11761, 11859. .12000,12520. .12897)) /note="ORF3"	
	/codon_start=1 /product="hypothetical protein" /db_xref="GI:13620168"	
	tr/translation="MSDPGQNFNESTSCWAPFPSPRAILLAMIDQAEDDGLDILISEM FAPPTNIPRHSADQSEQRSGLDGRLAARVGEPPDIDENITSPSAISRTSTVPPPY SAIPNPIFTSTLQSPNMTVDSSQMLPAVSDVTNIGPGQVETSGEANAPLITNNDB PYQVNAADILDLVDFDVPTEESVYLPIYETDVAERIDSLVPSNFEHEFVGDAVNVN IISNEGSEVSDSEDDLEDFEDEFEDDEAEAEADNDENDKADEDDNDGDGDEDE EDDEPDENEFEADHDDEADPDAAEESPKRKRPASSNMIGARTNTAAQVYIIQME DEDDPEDEGFWKRYGCKQVYKGNPNPSRYKCTNNDCNVKKHYERGAQDNFKLIVTSYDE IHNNPEPPACRITNSGPRNRSCTTTTQOQSYRDRIGFPAPASVITVPMEMPLAES TPYLDMSKYVRMTGLKLPNFSVFQNPGMTIRDEPMANVYNGSGIYGGIHRRLFAE GVNLM"	
intron	complement(10954. .11108) /note="ORF3"	
exon	/number=4 complement(11109. .11264) /note="ORF3"	
intron	/number=4 complement(11265. .11358) /note="ORF3"	
exon	/number=3 complement(11359. .11761) /note="ORF3"	
intron	/number=3 complement(11762. .11858) /note="ORF3"	
exon	/number=2 complement(11859. .12000) /note="ORF3"	
intron	/number=2 complement(12001. .12519) /note="ORF3"	

[illegible]

LOCUS	DEFINITION	SEQUENCE	13 FROM	PATENT	MO9846759.
LOCUS	8	687 bp	DNA		
LOCUS	A84084				
DEFINITION	Sequence 13 from Patent MO9846759.				
ACCESSION	A84084				
VERSION	A84084.1				
KEYWORDS	GI:6733223				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; kosidae; eurosids II;				

REFERENCE	Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 687)
TITLE	Theres.N.
JOURNAL	PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED ABSCISSION AREA FORMATION
FEATURES	Patent: NO 9846759-A 13 22-OCT-1998;
source	THERES NIKOLAUS (DE)
BASE COUNT	location/qualifiers
ORIGIN	1. 687
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
Query Match	12.7%; Score 218.8; DB 9; Length 687;
Best Local Similarity	59.2%; Pred. No. 9.8e-37;
Matches 392; Conservative	0; Mismatches 267; Indels 3; Gaps 1;
QY	707 cctgtctccacactcttcgcgaatcacgcggtgactggaatgaccttgatacccttcgtgaaca 766
DB	25 CCACCTCCATCTCTCCGCAATACCGATGCGGTCGAGATGTAACCGATTAACCGACT 84
QY	767 ggtgtcgttttagcttaaatctgtccatccattagggttgagtttcaattccatccctct 826
DB	85 GGAGACCGGTTTAACCCGGTTCGCTGACACTTTTAGGTCTCCAAATTCAGTTTCACACGCTA 144
QY	827 tataatccaataaataaccacgatacagatgagaatccttcctatatttccctccatgta 886
DB	145 GGTATCGTATGAGAAAGATGCTCCGCGACTTTTGTATACGATGTCGATGTTGTACTCTGTCA 204
QY	887 ctactccctgatatgaacctagctatacaactgltttcttaccctccaccgctttaa 946
DB	205 GCCGTATCAAGGAGAACCACTTCCCTCAATTTGTTCACTTCTCCACAAATATTATTAAAC 264
QY	947 gac---cgcgaaagttaagatcttttttgcataggttaagtcaatgaaccttaaat 1003
DB	265 GACGATGAGATATATATGTCGTCATCTTTGTGACGATCAAGACTTAACCTGTGAAATC 324
QY	1004 gttacaatcgcgagaaagaaacaaataacataccatccctcttllttlacaagaattcatc 1063
DB	325 GTTACAATGCGCAGAGAGAAAGACTATATCATGAGATCACTGCTTCTGAAATGATTTCT 384
QY	1064 gaggcgcttgattatatacaagctgltgttgattcaactggaagctacatctccaccgagt 1123
DB	385 GAGGCAAGTGCATCATATCATATGCGCATCTTTGATGCTGTGGAAGCGACGTTGCCGCAAT 444
QY	1124 agtcgaagagagatgaacagttgaacaagtgtggtttggagagagatgtttatatcgt 1183
DB	445 AGCCGAGAGAGACTTAACCTTAGAGAACGCTGCTGGTGAAGAAATTTGATGATTTTGG 504
QY	1184 ggcgaagagagagataaagaagaaagaaacatactgaagaagttagatcatcaggaagttag 1243
DB	505 GCGGCGGGAAGAACGAGAGAGAAAGCAAAAGACATTCGAGAGTTTGGGAAGAGATG 564
QY	1244 ttgagagagtttgagatttagtaagtgtgctttaaagccctttgcatlacaagaactaa 1303
DB	565 ATGAAGAGGTTTGGTTTGTCTTAACGTTCTATTGGAAGCTTTGCTTGGCTCAAGCTAAG 624
QY	1304 ctctcttttgagactcatatcatcctcttgaaagctatacaactcggaggttcgagtaatt 1363
DB	625 CTTCTCTTAAAGACTTCATTATTCCTTCAGAAAGGTTTAAATCTTCACTTCCCTTAACAATTC 684
QY	1364 tt 1365
DB	685 TT 686

```

VERSION      Y11336.1 GI:2339977
KEYWORDS     RGAI gene.
SOURCE       thale cress.
ORGANISM     Arabidopsis thaliana
              Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
              Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
              Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 2210)
              Truong, H.N., Caboche, M. and Daniel-Vedele, F.
              Sequence and characterization of two Arabidopsis thaliana cDNAs
              isolated by functional complementation of a yeast gln3 gdh1 mutant
              FEBS Lett. 410 (2-3), 213-218 (1997)
JOURNAL      97379310
MEDLINE      2. (bases 1 to 2210)
AUTHORS      Truong, H.N.
REFERENCE    Direct Submission
              Submitted (13-FEB-1997) H.N. Truong, INRA-Versailles, Laboratoire
              de Biologie Cellulaire, Route de Saint-Cyr, 78026 Versailles Cedex,
              FRANCE
FEATURES
  source      Location/Qualifiers
              1..2210
                /organism="Arabidopsis thaliana"
                /sub_species="(L). Heynh"
                /db_xref="taxon:3702"
                /dev_stage="two-leaf stage seedlings"
                /note="Landsberg erecta ecotype"
              132..1895
                /gene="RGAI"
                /gene="RGAI"
                /product="RGAI protein"
                /protein_id="CAA7217.1"
                /db_xref="GI:2339978"
                /db_xref="SPTREMBL:O23642"
                /translation="MKRDHROGRISNNSTSSSSSSISKDKMMVKKEEDGGGNNMD
                /ELAVLGYVRSSEKVAIKLEOLETMSNVPEGLSLATDVHYNPSELSMLDN
                MELSNPPLPSSNGIDVLPSPETICGPADYDKYIPGAITYOPALIDSSSSNS
                ONKRLKSCSPDSMTSTSTGTQIGVITVTTTAAESTBVIYLDSENG
                VRLVHLMACAEAIQONNLTAELVKQIGLAVSOAGMKRVATTFEALARKYRL
                SPONOVIDCLSTLOMHFETCPYLFKFAHFTANOAILFAFGKRVHVIDSMOGL
                OMPALOMALAREGPPPTFLNGIGPAPDNSDHLHEVCKLAQLAEALHVEFEYGF
                VANSIDIDAPSMELRSPDEAVANVSFEKHLKLRPGIEKVLGVKQICMLVACEG
                VDESEHETLSQWGNRSGSLPAHLGSHNARQASMLSVNSOGYVEESNCLM
                LGWHTRLITTSAMKLTAAH"
BASE COUNT   583 a 454 c 518 g 655 t
ORIGIN
Query Match      8.2%; Score 141.6; DB 13; Length 2210;
Best Local Similarity 52.8%; Pred. No. 3.3e-20;
Matches 453; Conservative 0; Mismatches 384; Indels 21; Gaps 6;

```

```

QY 854 gatgaagatcccttattatctccatctgactccctgataagaaacccagctatc 913
DB 1326 GATGTCGATGCTTCGATGCTT---GAGCTTAGACCGAGACGATACGAGAGCTGTGGGTG 1382
QY 914 aactggttttctactccaccgctttaaagaacgcgaaagtaagatttttg 973
DB 1383 AACTGCTTTTGACTACTACTAAGCTCTTAGGTGTCCCGGTGGAGATAGAAAGTTCTC 1442
QY 974 cataggttaagtaacatgaacccaaatctgtacaacgcgaggaaggaagaacatcat 1033
DB 1443 GCGGTGTGAACAGATTAACCGGTGATTTTCACGGGTGTGAGCAAGAAATGACACAT 1502
QY 1034 aaccatccctttttttaaagaattcagagcgctggattatatacagctgttt 1093
DB 1503 AACGACCGGTGTTTAAACCGGTTTACTGAAATGCTTAACTATATTATTCGACTGTGTT 1562
QY 1094 gattcactgaaagctacatctgaccacggttagtcgagagagatgacagtgcgaagtg 1153
DB 1563 GATTGCTGGAAG-----GAGTTCGGAATAGTCAAGAACAAAGTCATGCTG---AAGTT 1613
QY 1154 tggtttgagagagatgtttgatatcgttcgatatgagagagataaagaagaaga 1213
DB 1614 TACTTAGGGAACAGATTTGTAATCTGTGCTGTGAAGGCTTGACAGAGTCCAGAGA 1673
QY 1214 catgaagattatgataatgaagatgataatgagagagttgagattgaaatgttgc 1273
DB 1674 CACGAACGTTGATCAATAGGGAACCGGTTGTTGTCGCCGTTTACGCCGACAT 1733
QY 1274 ttaagccctttgcatatatacaagaagtaagctctcttgaagctatattatctctgaa 1333
DB 1734 CTGCGGTCTAACGCGTTTAAAGCAAGCAGATATCTTTGCTGTGTTAATGTCGCA 1793
QY 1334 ggcatacaactcggagtttcgagtaattctctcttctaagttggaataacacccct 1393
DB 1794 GGTATTCGTGTGAGAGATGATGATGTTTATGTTGGTGGTGCACACTGCCCATC 1853
QY 1394 ttctccatctcgtctgg 1411
DB 1854 ATTACCACTCCGCTTGG 1871

RESULT 10
LOCUS       ATY15194 2201 bp mRNA. PLN 01-NOV-1997
DEFINITION Arabidopsis thaliana mRNA for GRS protein.
ACCESSION   Y15194
VERSION     Y15194.1 GI:2569939
KEYWORDS    GRS protein.
SOURCE      thale cress.
ORGANISM    Arabidopsis thaliana
              Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
              Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
              Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 2201)
              Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J.,
              Murphy, G.P. and Harberd, N.P.
              The Arabidopsis GAI gene defines a signalling pathway that
              negatively regulates gibberellin responses
              Genes Dev. in press
              Harberd, N.P.
              Direct Submission
              Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular
              Genetics, Colney Lane, Norwich, NR4 7UJ, UK
FEATURES
  source      Location/Qualifiers
              1..2201
                /organism="Arabidopsis thaliana"
                /strain="Colombia"
                /db_xref="taxon:3702"
                /chromosome="2"
                /codon_start=1

```

```

/product="GRS protein"
/protein_id="CA75493.1"
/db_xref="GI:2569940"
/db_xref="SPTREMBL:O23725"
/translation="MKRDHOFQGRLSNHSSTSSSSSSISKDKMMVKKEEDGGNMD
ELLAVILYKVRSSMAEVALKLEQLTMSNVQDGLSHLADVHNPSELXMDLN
MSELNPPPLPSSNGIDPVLSPDPCGFPVLPENALYQFPAIDSSSSNN
ONRKLSKSPDSSWVTSNGTOIGVIGFTVTVTAAESTRSLVLDSENG
VRLVHLMACAEATOONNTLAEALVIOGALVSOAGMKRVAFTPEALARRIYEL
SPRONQIDHCLSDTLQMHFTYPCYLFKFAHTANQALILEAEGRKRVHIDPSNQG
OWPALQALALREGPPFRLTGIGPPADSDHLHEGCKLQALAEVHVEFEYRGR
VANSIADLDSMLELRSDPEAVANVSFELHKLIGRGQIEKYLKQVKKPIYFTV
VEQSNNGNPVFLDRFETSLHYSTLFDLSEGVNSQDKVSEYVIGLQICNLVACEG
PDVERHETLSOMGNRFGSSGLAPAHLGNSNAKQASMLSVFNSGGQVRVESNGCLM
LGMHTRPLITTSAMKLISTAAY"

```

```

BASE COUNT      576 a      450 c      525 g      650 t
ORIGIN
Query Match      8.1%; Score 140; DB 14; Length 2201;
Best Local Similarity 52.7%; Pred. No. 7.2e-20;
Matches 452; Conservative 0; Mismatches 385; Indels 21; Gaps 6;

```

```

Oy 560 ccttcataaaggttaccataaaccgataacagcgatttagaagcgataacggt 619
Db 1041 CCTATCTTAATTCGCTCCTTACGAGGACCAAGCATTCCTGCAAGCTTTTGAAGCT 1100
Oy 620 aatcacaagcaatccatcgtttagttagcacaataacagggggttcaatggccacg 679
Db 1101 AAGAAGAGAG---TACAGGTCAATGATTTTCGAGAACCAAGGCTCTCAAGGCGCTGG 1157
Oy 680 ttaatgaacagcactagcgtatcgtla---ccctgtcccaactcttcgaataccggtact 736
Db 1158 CTTATGCAAGCCTTGCCCTTGCAGAGAGAGGTCTCCCACTTTCCGGTTAACCGGAATT 1217
Oy 737 gg---aaatgaccttgaatacccttcgtagaaggttgatcgttagtaaatcttcac 793
Db 1218 GGTCCACGCGCGCCGGAATAATCTGATCATCTTCAAGAGTTGGTTGTAATTAGCTCAG 1277
Oy 794 tcataggttgagatttcaatcattcattcctcttatatagcacaataataacagatcac 853
Db 1278 CTTGCGGAGGCGGTTACAGTAAATTCGAAATCCGTGATTCGTTGCTAACGCTTAGCC 1337
Oy 854 gatgaagatcctcttatttccatcattgatacctccctgataaacccttagctatc 913
Db 1338 GATTCGATGCTTCATGCTTAGC---TTAGACCGAGCGATACGGAACCTTTGGGGTG 1394
Oy 914 aactgtgttttctacccacccgctttaaagaacgagcgttaagaattttttg 973
Db 1395 AACTGTGTTTGGACCTACATTAAGCTCTTAGTCTGCCGGTGGGATGAGAAAGTTCTC 1454
Oy 974 cataggttgaagtaacaaaccttaaatgttacaatcgcgagagaagaagaacaatcat 1033
Db 1455 GCGGTTGGAAACACATTAACCGGTTGATTTTCACGGTGGTGGACCAAGAAATCGAACCAT 1514
Oy 1034 aacctcctctcttttcaaaagattcgcaggttgattatatacagcttgitt 1093
Db 1515 AACGACCGGTTTCTTGACCGGTTTACTGAAATGTTACATTAATTAATTCGACTCTGTT 1574
Oy 1094 gatcactggaagcatttgcaccggtagtcgagagaagatgacagtttaacaacgtg 1153
Db 1575 GATTGCTGGGAAG-----GAGTTCGGAATAGTCAAGACAAAGTCATGCTC---AAGTT 1625
Oy 1154 tggtttgagagagagattgttgatatacgttcgtagaagaagaagaagaagaaga 1213
Db 1626 TACTTAGGAAACAGATTGTGATCTGGGCTTGTGAAGGCTCTGACAGAGTCGAGAGA 1685
Oy 1214 catgaaagtttagatcatcgtggaagtattgttgagaggttgcattagtaattgtct 1273
Db 1686 CACGAAACGTTGAGTCATGAGTGGGAAACCGGTTTGGTTCGTCGGTTTAAAGCCGACAT 1745
Oy 1274 ttaagcctttgcatatcaagaagtaagctcttttgagagctcatatcctctgaa 1333

```

```

Db 1746 CTTGGGCTAAGCGCTTAAAGCAAGCAGATGATCTTTCTGTGTTAATAGTGCCAA 1805
Oy 1334 gctataactcgtgaggttcgagtaattcttctcttagtgtagaataaacccct 1393
Db 1806 GGTATTCGTGTGAGAGAGATGATGATGTTGATGTTGGTGGCAGCATCTGCACATC 1865
Oy 1394 tctccatctcgtctg 1411
Db 1866 ATTACCACTCCGCTTGG 1883

RESULT 11
AC005560/c
LOCUS
DEFINITION
Arabidopsis thaliana chromosome II section 3 of 255 of the complete
sequence. Sequence from clones F10A8, F219.
ACCESSION
AC005560 AE002093
VERSION
AC005560.2 GI:6598472
KEYWORDS
HTG.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 95137)
Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
Fuji,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
Feildlyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Rensing,C.M.,
Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., Vanden,S.E., Umayam,L.,
Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
Venter,J.C.
Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana
Nature 402 (6763), 761-768 (1999)
JOURNAL
MEDLINE
20083487
PUBMED
10617197
REFERENCE
2 (bases 1 to 95137)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Dec 17, 1999 this sequence version replaced gi:3785968.
The sequence and annotation of chromosome 2 were merged from those
of the individual clones on this chromosome after removing
overlaps. For detailed information, please see the TIGR web site
(http://www.tigr.org/tdb/at/at.html).
COMMENT
Genes were identified by a combination of three methods: Gene
prediction programs including GRAL
(ftp://arthur.epm.ornl.gov/pub/xgral), GeneFinder (Phil Green,
University of Washington), Genscan (Chris Burge,
http://genomic.stanford.edu/GENSCAN/htm), and NePlantGene
(http://www.cbs.dtu.dk/services/NetGene2/), searches of the
complete sequence against a peptide database and plant EST
databases at TIGR, and manual curations based on those analyses.
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by two
or more gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats were
identified by RepeatMasker (Arjan Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Genes are
numbered from the top to bottom of the chromosome.

```

We thank the GSHR/Mashu/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards



and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: ateligr.org.

FEATURES  
source location/Qualifiers  
1. 95137  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="11"  
misc\_feature  
/note="Sequence from clone F10A8"  
301..>95137  
/note="Sequence from clone F219"  
341..>997  
/gene="At2g01390"  
/note="F219.1; predicted by genscan"  
341..997  
/gene="At2g01390"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="AAC67316.1"  
/db\_xref="GI:3785969"  
/translation="MKONLAVADILINOMRDNRIKIDFVVSALITNCDRCRREGAS  
LAVDYLEMGIHLKKSAYLALIGNFSLNEPLKVIIEVKEWKAQHSIGCYGAVLILH  
RIGCFGRPLADVDPLDPDOKGVAAATLADVYISAGSPKAKRIIREMEREIMP  
SLCTYVLLSGLEKTSIDPQKEVALLRKEKSLVASARENVHEDKICDILFATNLL  
"  
repeat\_region  
complement(1085..1127)  
/rpt\_family="(TAAA)n"  
join(<1175..1294,1620..>1940)  
/gene="At2g01400"  
<1175..>1940  
/gene="At2g01400"  
/note="F219.2; predicted by genscan"  
join(1175..1294,1620..1940)  
/gene="At2g01400"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="AAC67317.1"  
/db\_xref="GI:3785970"  
/translation="MARSCSITMTVPVILSLSPVTGESPKISRRVILATISGSPOLLE  
AKIKEPKLOVLDKSDYRRNTMTYLAAGPLGINFLNGEAARVGRKREKKALEK  
LRARAKSEEPNNKSGNOKIRKEKEVFPILPPPLVEANILQ"  
mRNA  
complement(2069..>3232)  
/gene="At2g01410"  
complement(<2069..>3232)  
/gene="At2g01410"  
/note="F219.3; predicted by genefinder"  
complement(2069..>3232)  
/gene="At2g01410"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="AAC67318.1"  
/db\_xref="GI:3785971"  
/translation="MKANSSSSSRYSRFFVPIFLFAVFLGILIPSSADNRHVINP  
KSPGLIPGLIWDPRDOHETVIGSLHSRTISVAGVETLISDLDPENSTIGLAV  
DSTNRRLACIOSLPPLPPFSALASYSIDRSGRVETLSPLDIPGDEEDLADVANDV  
AVDFGNAYVYVNSAKNFIWVADROGAATFSKSPLENSQPAADADASPDGCGNGIV  
YISGCVLLVQVSNTEKRVKVDKSGNARVLVINDLLAADETRRRDGYVYVSQK  
LMLKSDSSEGVYDEIDLDIEGFPAVAVAGDRLYVLYGVMGIMSSKKEGS  
AREWFGIEVMSKEGEGDKIMLYLIGFGVAYCFMRFQMKKLITNDDKKT"  
mRNA  
complement(join(6967..7404,7509..7585,7680..7837,  
8215..8300,8440..8707,8865..>10059))  
/gene="At2g01420"

gene complement(6967..>10059)  
/gene="At2g01420"  
/note="F219.4"  
complement(join(7338..7404,7509..7585,7680..7837,  
8215..8300,8440..8707,8865..10059))  
/gene="At2g01420"  
CDS  
/codon\_start=1  
/product="putative auxin transport protein"  
/protein\_id="AAC67319.1"  
/db\_xref="GI:3785972"  
/translation="MITWHDLYTLAVVPLVYVAMILAYGSVQWKKIFSDQSGINR  
EVAIRAVPLSPHETSTNDPYAMNFRFVAADTLQKIMLYVLLALAMILNKGSLKEMI  
TFESITLPTNLTIGITPILIAMGYVAGSLVQVLOCITWYLLILEEYRGAKLL  
IMEQPEFGASTVSEFKVESDYSDLDHSPFYVSKPEPGRGLSNRPADLYSVSSRGPT  
TPRPSNLGAEIYLSSTPRGSENFHSDPFPYVSKPEPGRGLSNRPADLYSVSSRGPT  
PPSPNEENNAVYGFYNNNTSSVPAGSYFPAPEPSTGCVSTKPKIKEMOOL  
QKDSKASHDAKELHMFVWSSSASPVSDFGGAGDNVATQSDQGAKEITMYSDP  
RKSNAAGGDDIGLDGSGEREIEKAAGLNKGMSNTAELEAGGDSGNGTTHP  
PISVMTRLIIVWRLKINPNTYSLIGLWALVAVRWVAMRKLQOSTISLSDXG  
LCAWPSIGLPMALQPKIYAGNSVATFAMAVRFTGPAIMAVVGAIGLGLDLRLA  
IVQALPOSTVPEVFAKEYNHPTLISGVIFGMLALPILYVILLGL"  
14152..14264  
/rpt\_family="POLY A"  
complement(join(<14658..14916,16258..16337,16458..>16607))  
/gene="At2g01430"  
complement(<14658..>16607)  
/gene="At2g01430"  
/note="F219.5"  
complement(join(14658..14916,16258..16337,16458..16607))  
/gene="At2g01430"  
/codon\_start=1  
/product="putative homeodomain transcription factor"  
/protein\_id="AAC67320.1"  
/db\_xref="GI:3785973"  
/translation="MNRIPSEEDGDEEFSHDGSAAPPKRLITRREOSRLLEDSEFRQ  
NHTLNPQKQVYLAKHLMLRPROJEVNRFRARRSKIKOTEMCEVYLRMGSLTEENH  
RLHREVELRAMVGGPTVNASASLIMCRCEVTPAASPAAVVPVPAKKFPPOER  
DR"  
complement(join(<20954..21559,21680..21763,21869..22027,  
22234..22377,22526..22606,22742..22813,23017..23250,  
23352..23441,23559..23640,23906..24069,24308..24542,  
24746..24858,24942..>25415))  
/gene="At2g01440"  
complement(<20954..>25415)  
/gene="At2g01440"  
/note="F219.6"  
complement(join(20954..21559,21680..21763,21869..22027,  
22234..22377,22526..22606,22742..22813,23017..23250,  
23352..23441,23559..23640,23906..24069,24308..24542,  
24746..24858,24942..25415))  
/gene="At2g01440"  
/codon\_start=1  
/product="putative ATP-dependent DNA helicase (RECG)"  
/protein\_id="AAC67321.1"  
/db\_xref="GI:3785974"  
/translation="MECDNVDDFTDKRSDQVKKDLVACKRPPSIILGDSRPEVLS  
NSKSGESSITLPTDINSFLTPMHGGMPEPDMISLTSFCEPLLQNDSDSPRD  
IIDDSSFTSTKATSEVATSDVFAADRLATSIDMPCLSKHSNQLDSCGHTK  
KLHHPFRTYADIQNAQVYDIDGOYLIFGVLSKGRASSSPSEVLYVSGSGR  
DKTPADHFEITRYVGLPSLNDAYVGHPEPTLDEADLRKRLTFDEFYLIQARLQ  
PAMDPIDREITRYVGLPSLNDAYVGHPEPTLDEADLRKRLTFDEFYLIQARLQ  
SLRADHFEITRYVGLPSLNDAYVGHPEPTLDEADLRKRLTFDEFYLIQARLQ  
MLQSGTKLEKVDLKEFKRPVLSNVYLEESTLTKSLKLPSTLPSQSAVSELI  
WOLKRPVNNRLLOGDVGKTVAVFLACMEVYISGYOAAVMAPELTALIOHYQCD  
LLENMGDOSATSPITIGTSHLAKETLSYSLRLVAVVDEQGRVIGDIOGFNKGK  
TSMISGSSSDSDTSKADLSMAPIVAMLSATPDISLALYDGLISLTOIGMLPR  
IPVETHIFEENGTGIEKYSMLDELKSGGKYVYVYIIDSQDLPOLRAASALEIY

Query Match 8.1%; Score 140; DB 12; Length 95137;  
Best Local Similarity 52.7%; Pred. No. 9, 5e-20;  
Matches 452; Conservative 0; Mismatches 385; Indels 21; Gaps 6;  
Oy 560 ccttcaagaagttactcaataaccgcgaatcaagcgagatttaagaagcgttaacggt 619





QY 974 catagggttaagtaagaacccttaaaatgttacaatcgcgcgagaagaagaacatcat 1033  
 Db 1348 GGTGGGTGAATCAAGATTAACCGAGATTTTCACTGGTTGAGCAGGAATCAACCAT 1407  
 QY 1034 aaccatcctcttttttacaagaattatcagaagcgttgattatatacagaatgtgtt 1093  
 Db 1408 AATAGTCCGATTTCCTTAGATCGGTTTACTGATGCTGTTCATTTTACTGACGTTGTT 1467  
 QY 1094 gattcaacggaagctacatctgccaccggttagtcgagagaagatgacagtgtgaacaagt 1153  
 Db 1468 GACTCGTTGGAAGGT-----GTACCGAGTGTCAAGACAAGGTCA--TGTCGAGAGTT 1518  
 QY 1154 tggtttggagagagatgtgtatcgttgatcgtgagagaagataaagaagaagaaga 1213  
 Db 1519 TACTGGGTAACACATGCTGCAACGTTGCTGTGTTGATGACACTGACCGAGTTGAGGTT 1578  
 QY 1214 catgaaggttagatcatcaggaagtaagtgtgagaggtgtgagattagtaagtgtgt 1273  
 Db 1579 CATGAAGCTTGATGTCAGTGGAGAACCGGTCGCGTCTGCTGCGGTTGCGGCTGCACAT 1638  
 QY 1274 ttaagccctttgcatatcacacaagtaagctctcttttggagactcatatccttctgaa 1333  
 Db 1639 ATGTGTTCAAAATGCGTTTAAGCAAGAGACTATGCTTTGGCTCTGTTCACACGGCGGTGAG 1698  
 QY 1334 ggtctaacctcgaagtttcgagtaattcttctctttagtggtaaatcaaccctt 1393  
 Db 1699 GGTATTCGGGTGAGAGAGTGAACGGCTGTCTCATGTTGGTGGCACACACGACCGCTC 1758  
 QY 1394 ttctcaatcgtctgtg 1411  
 Db 1759 ATAGCCACTCGGCTTGG 1776

RESULT 13  
 LOCUS A64697 1964 bp DNA PAT 29-MAR-1999  
 DEFINITION Sequence 1 from Patent WO9729123.  
 ACCESSION A64697  
 VERSION A64697.1 GI:4530762  
 KEYWORDS  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 1964)  
 AUTHORS Harberd,N.P., Peng,J., Carol,P. and Richards,D.E.  
 TITLE NUCLEIC ACID ENCODING GAI GENE OF ARABIDOPSIS THALIANA  
 JOURNAL INNES JOHN CENTRE INNOV LTD (GB)  
 COMMENT Patent: WO 9729123-A 1 14-AUG-1997;  
 FEATURES Other Publication AU 1799697 19970828.  
 FEATURES Location/Qualifiers  
 source 1..1964  
 /organism="Arabidopsis thaliana"  
 /strain="LANDSBERG ERECTA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /map="20.9"  
 /haplotype="GAI"  
 BASE COUNT 489 a 426 c 474 g 575 t  
 ORIGIN

Query Match 7.7%; Score 133.6; DB 9; Length 1964;  
 Best Local Similarity 52.2%; Pred. No. 1.7e-18;  
 Matches 448; Conservative 0; Mismatches 389; Indels 21; Gaps 6;  
 QY 560 ccttataaggttactaatcaacgcgttaatacagcattttagaagcgatgaaggt 619  
 Db 947 CCTTATCTCAAGTTCCTCACTTCAACGGAGATCAAGCATTTCTGAACTTTTCAAGGG 1006  
 QY 620 aatcataagcaatcacatcgtttgatttcgacataatcaacgaggttcaatggccacgg 679

Db 1007 AAGAAAGAG--TTCATGTCATTGATTTCTTATGATCAAGGTCTTCAATGGCCGGCG 1063  
 QY 680 ttaatgaacacactagctgcatcgttaccctgtcccaactt---cgatcacccgttact 736  
 Db 1064 CTTATGACAGGCTCTTGCCCTTCGACCTGCTGTGCTCTCTGTTTCCGTTACCGGAATT 1123  
 QY 737 gg---aaatgaccttgatacccttcgttagaacaagtgatcgtttagtaaatltgtcac 793  
 Db 1124 GGTCCACCGCGCACCGAATTAATTTGCATTATCTCATGATGAAGTTGGGTGTGAAGTGGCTCAT 1183  
 QY 794 tcaatgaggttagatcttcaatccatccctcttataatagcaataaataaccagatcac 853  
 Db 1184 TTAGCTGAGGCCA---TTCACGTTGAGTTTGTAGTACAGAGATTGTGGCTTACACTTTA 1240  
 QY 854 gatgaagatcctcttataatcttccatccttactaccctccctgttgaacacctgatac 913  
 Db 1241 GCTGATCTTGATGCTTCGATGCTTGAGCTTACACCAAGTGAAGATTGAATCTGTGGGTT 1300  
 QY 914 aactgttttctacatccacgcgctttaaagaacgcgaaagttaaagattttttg 973  
 Db 1301 AACTCTGTTTTCAGAGCTTCAACAAGCTTGGAGCAGCCTGGCGATGATGAAGTCTTT 1360  
 QY 974 catagggttaagtaagaaccctaaatgttacaatcgcgcgagaagaagaacatcat 1033  
 Db 1361 GGTGGGTGAATCAAGATTAACCGAGATTTTCACTGTGTTGACAGGAATCAACCAT 1420  
 QY 1034 aaccatcctcttttttacaagaattatcagaagcgttgattatatacagctgtgtt 1093  
 Db 1421 AATAGTCCGATTTCCTTAGATCGGTTTACTGAGTGTCTCATTAATACGACGTTGTT 1480  
 QY 1094 gattcaacggaagctacatctgccaccggttagtcgagagaagatgacagtgtgaacaagt 1153  
 Db 1481 GACTCGTTGGAAGGT-----GTACCGAGTGTCAAGACAAGGTCA--TGTCGAGAGTT 1531  
 QY 1154 tggtttggagagagatgtgtatcgttcgttcgagtaagaagataaagaagaagaaga 1213  
 Db 1532 TACTTGGGTAAACAGATCTGCAACGTTGCTGTGATGATGACACTGACGAGTTGAGCTGT 1591  
 QY 1214 catgaaggttagatcatcaggaagtaagtgtgagaggtgtgattagtaagtgtgt 1273  
 Db 1592 CATGAAGCTTGATGTCAGTGGAGAACCGGTTGCGTCTGCTGCTGCTGCGGCTGCACAT 1651  
 QY 1274 ttaagccctttgcatatcacacaagctctcttttggagactcatatccttctgaa 1333  
 Db 1652 ATGTGTTCAAAATGCGTTTAAGCAAGAGACTATGCTTTGGCTCTGTTCACACGGCGGTGAG 1711  
 QY 1334 ggtctaacctcgaagtttcgagtaattcttctctttagtggtaaatcaaccctt 1393  
 Db 1712 GGTATTCGGGTGAGAGAGTGAACGGCTGTCTCATGTTGGTGGCACACACGACCGCTC 1771  
 QY 1394 ttctcaatcgtctgtg 1411  
 Db 1772 ATAGCCACTCGGCTTGG 1789

RESULT 14  
 LOCUS ATY15193 1964 bp DNA PLN 01-NOV-1997  
 DEFINITION Arabidopsis thaliana GAI gene.  
 ACCESSION Y15193  
 VERSION Y15193.1 GI:2569937  
 KEYWORDS GAI gene.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 1964)  
 AUTHORS Peng,J., Carol,P., Richards,D.E., King,K.E., Cowling,R.J.,  
 TITLE The Arabidopsis GAI gene defines a signalling pathway that  
 JOURNAL negatively regulates gibberellin responses  
 Genes Dev. In press

REFERENCE	2 (bases 1 to 1964)
AUTHORS	Harberd, N.P.
TITLE	Direct Submission
JOURNAL	Submitted (22-Oct-1997) N.P. Harberd, John Innes Centre, Molecular Genetics, Colney Lane, Norwich, NR4 7JL, UK
FEATURES	Location/Qualifiers
source	1 1964

BASE COUNT ORIGIN	489 a	426 c	474 g	575 t
----------------------	-------	-------	-------	-------

Qy	1094	gattcactcggaaagctacattcgtccaccgggtagtcgagagaaatgatacagttgaaacaagtg	1153
Db	1481	GACTCGTTGGAAAGT-----GTACCGAGTGGGCAAGACAAGCTCA--TGTGGAGGTT	1531
Qy	1154	tgcttctggagagagatctgtgtatcgccttcgcatgagagagatacaagaagaagaaga	1213
Db	1532	TACTTTGGGTAAACAGACACTCCAAACGTTGTGCGCTTGTGATGAGACCTGACCGAGTTGAGCT	1591
Qy	1214	catgaagagttatgatcatcatcgtggaagatattgttgaggagtttggaattagtaattgtct	1273
Db	1592	CATGAACGTTGAGTACGTAGTGGAGGAACCGGTTGCGGTCTGCTGGGTTTCCGGCTGCACAT	1651
Qy	1274	ttaagcccttctgcatactacaagaagtcctctctttgagaactcatatcctctcgtaa	1333
Db	1652	ATTGTTTCGATAGCTTAAAGCAAGAGATATCCTTTGGCTGTGTCACACGGCGGTG	1711
Qy	1334	ggcattcaacccgagtttcgagtaattcttctctcttgtagtttggaatatcaaacccgt	1393
Db	1712	GGTATATGGGGGGGAGAGAGTGAACGGCGTCTTCATATGGTGGTGGCACACAGACCGCTC	1771
Qy	1394	ttcccatctgccttg	1411
Db	1772	ATTAGCACCTCGGCTTGG	1789

RESULT	15		
AC006917			
LOCUS	AC006917	132699 bp	DNA
DEFINITION	Genomic sequence for <i>Arabidopsis thaliana</i> BAC F1086 from chromosome		
ACCESSION	AC006917		
VERSION	AC006917.6	GI:4757662	
KEYWORDS	HTG.		
SOURCE	thale cress.		

Oy 1034 aaccatccctctttttacaagaattcatcgagcgttgatatattacagcgtgatt 1093  
||| ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1421 AATAGTCGATTCTTACATCGCATTACTGCATTAATTAATCAATCGACGTGGTT 1480

**TITLE** Direct Submission  
**JOURNAL** Submitted (28-JUN-2000) *Arabidopsis thaliana* genome center,  
Department of Biology, University of Pennsylvania, 38th and  
Hamilton Walk, Philadelphia, PA 19104-6018, USA  
**COMMENT** On May 7, 1999 this sequence version replaced g1:4731042.

FEATURES		Location/Qualifiers
SOURCE		1..132699
CDS		/organism="Arabidopsis thaliana"
		/db_xref="taxon:3702"
		/chromosome="1"
		complement(join(144..278,360..491,576..673,814..1435,1749..1867,1161..2310,2441..2565,2698..2742,2837..2902,2979..3031,3117..3234,3314..3415,3498..3655,3741..3818,3894..3985,4054..4104,4138..4420)) BA89487.1"
		/note="similar to Na/H antiporter dbj BA89487.1"
		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.1"
		/protein_id="AAF79251.1"
		/db_xref="GI:8778242"
CDS		/translation="MTSTIGALPYKSPPEKAIASSSSYSAENDSPVDVIAFAGTSLVLTGACRYLNGTRPVYVAVLVLTGISTSLGICLTKLISYSISIMKILTEKYTKNLGKLGIRINGINDPDLAVFLPVLLFESSFSMDVHQIKRMQKMLLAGPVILSTFCIGALIKLTPPYMMDKTSLILGLGARDPVAVALLKEIGASKMTLLIDGESILNDGYSVYVOLFEFKVMGHNDSGSIKFLVQNSFGAVGIGLAFGLASVWLKFIENDTVAQITVLSAFVAYTAQEMAGVSGILTVMLIGMFFFAAFARAFKSGDSHILHFM EAAITAVTLPMLSGVITIASVSGOTISYKVSQFILLRPDFFILPLEDYKGGN SMSLELLYVQSLKRCVAVGLYPLICRSYGLDMKRESILTMGLRAVLOEFLDEL GIVFLVNGSTQQLHLMDPLATKRIILEYTRFEMKRAKALAPENIDDEL GSADWPVIRHISILKDEGRQVPHDYEAGSLDPNIMDIRIFELNGISOCVYSS HCOISVVKVAIDSRILFVSGVQAAWEMDDGRIQTCTANVLQSDVADLVSTSS LDMWGLEPVRHFPNYKFLQSKIIPHLVTHLIVERLESACYISSAFLRHRIARQQLHIFGLNSIASTVINESEGEAKQFLEDVRSFQVLSVLTQVTHYVNLHNG YTKNEKYGILLGKREKSHLHDVQSDLKKLRHPPSLKLPVVDLITSNPLKDRSSF RSLAIGEDA"
		join(5650..5801,6081..6121,6230..6819)
		/note="unknown protein; similar to EST emb F14394.1"
		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.2"
		/db_xref="GI:8778207"
		/translation="MRTPTTILLGAILFSGAGVVRSDASPHRYKEDVPLVANKY GPHNPESEYRFDLPFCIPESYKEKKMLGVLNGLDPLVASPYKLNFRDESESYC NKLTSKEVQKRAVEKDYFYOMTIDLPIMGFIGNYDKDKSDSPSKFYIKHIQ FEILYNKDRVIEISARMDPHSLVDLTDEKEVAEFYIVKMETETPEKRMKESMS SLPHLHEIHWFESINSCTVLLTDEKFLIIMRYLKNDEM"
		join(6881..7029,7140..7296,7392..7876,7961..8180)
		/note="similar to transmembrane 9 superfamily member 2 g1 4736874; similar to dbj av442489.1, gb A1996021.1, emb F14395.1"
CDS		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.3"
		/protein_id="AAF79217.1"
		/db_xref="GI:8778208"
		/translation="MKRYADEEADDOERQGWYIHGDVFRFPYHNSLFAASLGS G TQLPITLTFEMALGVFPYKNGALFTLAVLYALTSIAGTASASYCOLEGSS VRNLLTGCLGCLPFLFCFLNVALTVAIALPFGTIVVILWTLVWTSYVILVIG GIAGKNSAEFOACRTTKYPRLEIPWRSALPQMAAGLPFSAITELYIPAS VMGRITYISILFIVFLIILIVAFITVALTFOLEADHQMWSFGCGGTGLFTI YACLYYYVANSMSGMQTSFFPGYMACICGFYMLGTVGRALLVLRHLYRSIK GE"
		8842..9714
		/note="hypothetical protein"
		/codon_start=1
		/evidence=not_experimental
CDS		/product="F10B6.4"
		/protein_id="AAF79218.1"
		/db_xref="GI:8778209"
		/translation="WDISREYDDFKETIDHSLGLPISMDVLOKRLTAESQRRLNE /translaltion="WDISREYDDFKETIDHSLGLPISMDVLOKRLTAESQRRLNE OYLSVLRKEKETYIDVYRSEASMNQSLKGFVEENQKSGSEREDLVNCKKWEKOC FLYHODRESLMEFENDEDERAREASRRELEBEYKMSDEIKRSIESEDCIYDLSL ASFVAKDSISIGRIFLEANSBEDCCCTLKKMDQLKPTTKQKVSIVSWYKREKX ECLIMLAKAEQVELVSEONRELDRENKFLROCSARSHGSKFKRKSIRKMSSP IKRRIEISSOERID"
		11067..11906
		/note="unknown protein; similar to EST gb A1995472.1"
CDS		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.5"
		/protein_id="AAF79219.1"
		/db_xref="GI:8778210"
		/translation="MDDDFGRNMGYEPAAATFGKNGIGQLMSTIDRNTKPLPGRDP /translaltion="MDDDFGRNMGYEPAAATFGKNGIGQLMSTIDRNTKPLPGRDP NLMKPGNSHQPPIHMSYNMWINOQKRFPMPLPVLTATPYNVNLPEVSSASQO MNLHHHQTEENYVKKLEELIVQTKRKRNKAGSTPKAKPRKDEKSNMNNNT NVRVAKKSDVLYINGVMSDLSGLPVLICGTGAPQOCYRMGGGMOACCTNIS MHP NSTKRRGARIISGRKMSQCAFKVLKSLASDFNFGNPLDKSHWARHGTRKFV TIR"
		complement(13762..14781)
		/note="hypothetical protein"
		/codon_start=1
		/evidence=not_experimental
CDS		/product="F10B6.6"
		/protein_id="AAF79250.1"
		/db_xref="GI:8778241"
		/translation="MKLMKRAVVLKDGPSLIADDLTPAAVYKATSHDELSTDESA /translaltion="MKLMKRAVVLKDGPSLIADDLTPAAVYKATSHDELSTDESA OPIYRHVLSSESSSLKPLVSLISSVYKTRSMVAALKGLMHPFLCKSVASIGRL PBDLSFPEGNSRIMSKSGGNLFPVAVFALDRSLIFHDGNRHRIEESVLTLY IIRKQQLVDSLIRIKPIGENMMLPINEAMENYSEIMELGYMICKRIAEVLPNVHS KIKGTEDLAKIYAKSMKQGEIKKYFEPCDKDGVNAQEIENFVAIPEADVILHDE LVRTAMESSESAERTETAEEREEBEIEIKLSDLTLDHNEBPASPPRVVVVDIP DLISE"
		16027..16533
		/note="hypothetical protein"
		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.7"
		/protein_id="AAF79220.1"
		/db_xref="GI:8778211"
CDS		/translation="MSTCYVRECRNNAATKGSYALIDGREYOPSNGDLVACGCH /translaltion="MSTCYVRECRNNAATKGSYALIDGREYOPSNGDLVACGCH RSYHRIDVSSPQINHTREFPPTSRLRYKQALRLKWTAEERNEEDDTSTEEK MLYORRRKSKFTAEQREAKDYAAKLGWLMDKRLREIEIVFCBGIGVRYHKTWV NNNKRFYH"
		complement(join(17635..17815,17901..18019,18141..18194,18300..18406,18547..18611,18742..18950))
		/note="hypothetical protein"
		/codon_start=1
		/evidence=not_experimental
		/product="F10B6.8"
		/protein_id="AAF79249.1"
		/db_xref="GI:8778240"
		/translation="MEFKAYFLIYVFPRLCEPVIYLFQMICLKKRQIFLFW /translaltion="MEFKAYFLIYVFPRLCEPVIYLFQMICLKKRQIFLFW IPEBLHLOQPLFKWITFWLIRIRMWQDKOVCSSTTKGDSLVSAGAGALGFNE IVRPEYKSLGESFSLCSFKESKRYIRGOALKVGSILTFVEGAVRAGANIMIQSK EOSLIVSESSSFDENVNNKSSOSELCTILAKIFGSIANAIVGVGYARVLLPFV WKKDLGDSSSQNGSSDSBDTIDSD"
		complement(join(20071..20409,20499..20573,20648..20721,20800..20887,20957..21108,21188..21356,21448..21906,22006..22275,22367..22570,22840..22917))
		/note="hypothetical protein"
		/codon_start=1
CDS		/evidence=not_experimental
		/product="F10B6.9"
		/protein_id="AAF79248.1"
		/db_xref="GI:8778239"
		/translation="MLEIESPTSLCFRTNTTCAALLRELOKIWDIGESDAEKRMAL /translaltion="MLEIESPTSLCFRTNTTCAALLRELOKIWDIGESDAEKRMAL ELERKCEIYRRKYDEANASKAOLHOSLVSIEAEISLALGVFNHSPMKAKESGK SLKEKLAIVRPMLEDLRLQKDERMKQVDIRKAOIEKSGEISGSPQLNTMVGSLAL DEODITLTKRLNEOTHLRSLOKESDPLNVLVGVNIVVHLLCGVLDVQOYVEVAP SLHRTDHOSTNISDDTLDGJLHMHILKLTERTSRPYKQYKHCNMLILEVTDYHDA LAFSPSLQKDVAGSLFELMNLMDTSQDEETKTRASYVYRSESDIPEPNLISSET IEQVSAEYDCFNKLKASRMKELVAKRKREIELNLCRLAHIDADTISLEKTAALDSGT KRLVNPSELNTNIELHINKIKEEASHREILDIRDLRMLEDEKRSFLVDGVLVSLDEY AGRGCHVNLKHAERARITVNIKIPSWNDLILKTLTMDDEKRSFLVDGVLVSLDEY KILTRQAOBEERKRYRDOKKMODLILKRESIYSGKPPRRSNSRYKTINGNDASVPP TPRRSAGATNNDIMTTPRYSYSHRONYEFKEVRLSTAPLNFVAIPEKDSVSTYISV GSESDSPLYN"
		join(26661..26948,27454..27610,27773..27814,27911..28038,28342..28470,28565..28632,28747..28965)
		/note="similar to acid phosphatase type 5 emb CAB63938.1;

```
similar to EST gb|N96048.1"
/codon_start-1
/evidence-not_experimental
/product="F10B6.10"
/protein_id="AAE79221.1"
/db_xref="GI:8778212"
/translation="MTYIYRDYKRTTKSTIPFLIFLFCPSNLSMATLKHKPNLVVY
VYMLIIFSSHSSTAEIRLRLOPSKTDGTVSFLVIGDMGRGSYNQSOVALQMGELGE
```

Search completed: June 30, 2001, 05:28:45  
Job time: 4329 sec

```
Query Match 7.7%; Score 133.6; DB 12; Length 132699;
Best Local Similarity 52.2%; Pred. No. 2.3e-18;
Matches 448; Conservative 0; Mismatches 389; Indels 21; Gaps 6;

QY 560 ccttcataaggttaccataaaccgtaatacaagcattttagaagcattacggt 619
Db 118136 CCTATCTCAAGTTCGCTCAGCTTCAAGCGATCAAGCGATTCTCGAAGCTTTTCAAGG 118195

QY 620 aatcatcaagcaatccacatcggtgatttcgacataatcaacgaggttcaatgaccacg 679
Db 118196 AAGAAAGAG--TTCATGTCATGATTCTCTATGAGTCAGAGCTTCATAGGCGCGCG 118252

QY 680 ttaatgcaagcactagctgattcgcttaaccctgctccactct---cgaatcacggtact 736
Db 118253 CTATGCAAGGCTCTGGCGCTTCGACCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 118312

QY 737 gg---aaatgacctgatacccttcgtagaacaggtgactgtagctaaattgctcac 793
Db 118313 GGTCACCGCGACCGGATATTGATTTATCTTATGATGATGGGTGTAAGCTGGCTCAT 118372

QY 794 tcattaggttgatgattcaattccatccctcttatatagcaataataaccagatcac 853
Db 118373 TTACGTGAGGCGA--TTCACGTGAGTTTGATGACAGAGATTGTGCTAACACTTTA 118429

QY 854 gatgaagatcctctatattcctccatctgtaactcctccatgaaacccctagctatc 913
Db 118430 GCTGATCTTGATGCTTGATGCTTGAGCTTGACCAAGTGAATGATGATGCTGCGGTT 118489

QY 914 aacgtgttttctaccctccacgcttttaaaagaccgcaaaagttaagatttttgg 973
Db 118490 AACTCTGTTTTCGAGCTTCAACAGCTCTTGGAGACACTGCTGCGATCGATAGGTTCTT 118549

QY 974 cataggglttaagtaacctaanaattgtltaacatcgagaaagaaagcaaatcat 1033
Db 118550 GGTGTGATGATCAAGATTAAACCGAGATTTCACCTGTTGAGACAGAAATCGAACCAT 118609

QY 1034 aaccatcctcttttttaacaagattcagagcggttgattattatacagctggttt 1093
Db 118610 AATAGTCCGATTTCTTATGATCGGTTTACTGAGTCGTCATTTACTCGACGTTGTTT 118669

QY 1094 gattcactggaagctacatcgccacgggttagtcgagagagagatgacaggttgaacaagt 1153
Db 118670 GACTCGTTGGAAGT-----GTACCGAGTGGTCAACAGACAGTCA---TGTCCGAGGTT 118720

QY 1154 tggtttggagagagatgttgatatacgttcgaltggaagagagataaaagaaagaaaga 1213
Db 118721 TACTTGGGTAAACAGACTGCAACGTTGTGCTTGTGATGACCGTGAAGTTGAGCGT 118780

QY 1214 catgaaggtttagatcatcgtggaaggtatgttgagaggttgatgattagtaagtgtgct 1273
Db 118781 CATGAAGGTTGATCACTGAGGAGAACCGGTTCCGGGTTGCGGCTGCACAT 118840

QY 1274 ttaagcccttttgatltacacaaagtaagcttcttttgagactcatatcctcttgaa 1333.
Db 118841 ATTGCTTCGATGCGTTTAAACGACCGAGTATGCTTTTGGCTCTGTTCAACGCGCGTGA 118900

QY 1334 ggcatacaactcggaggttcgagtaattcttctcttaggttgcaaaatcaaccctt 1393
Db 118901 GGTATGCGGTGAGAGAGAGAGCGGCTGTCTCATGTTGGGTGACACACGACCGCTC 118960

QY 1394 ttcccatctcgctctgg 1411
Db 118961 ATAGCCACCTCGGCTGG 118978
```

---